(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 899 332 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.03.1999 Bulletin 1999/09

(21) Application number: **98301170.1**

(22) Date of filing: **17.02.1998**

(51) Int Cl.[6]: **C12N 15/12**, C07K 14/705, A61K 38/17, C12Q 1/68, G01N 33/68, G01N 33/50, C07K 16/28

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **15.08.1997 US 55895 P**
27.10.1997 US 962922

(71) Applicant: **SMITHKLINE BEECHAM CORPORATION**
Philadelphia Pennsylvania 19103 (US)

(72) Inventor: **Ellis, Catherine E.**
King of Prussia, Pennsylvania 19406 (US)

(74) Representative: **Crump, Julian Richard John et al**
FJ Cleveland,
40-43 Chancery Lane
London WC2A 1JQ (GB)

(54) The G-protein coupled receptor HFIAO41

(57) HFIAO41 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing HFIAO41 polypeptides and polynucleotides in the design of protocols for the treatment of infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias. such as Huntington's disease or Gilles dela Tourett's syndrome, among others and diagnostic assays for such conditions.

EP 0 899 332 A2

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] This application claims the benefit of U.S. Provisional Application No. 60/055,895, filed August 15, 1997.

FIELD OF INVENTION

[0002] This invention relates to newly identified polynucleotides, polypeptides encoded by them and to the use of such polynucleotides and polypeptides, and to their production. More particularly, the polynucleotides and polypeptides of the present invention relate to G-protein coupled receptor family, hereinafter referred to as HFIAO41. The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

BACKGROUND OF THE INVENTION

[0003] It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, Nature, 1991, 351:353-354). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., et al., Proc. Natl Acad. Sci., USA, 1987, 84:46-50; Kobilka, B.K., et al., Science, 1987, 238: 650-656; Bunzow, J.R., et al., Nature, 1988, 336:783-787), G-proteins themselves, effector proteins, e.g., phospholipase C, adenyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M.I., et al., Science, 1991, 252:802-8).

[0004] For example, in one form of signal transduction, the effect of hormone binding is activation of the enzyme, adenylate cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide, GTP. GTP also influences hormone binding. A G-protein connects the hormone receptor to adenylate cyclase. G-protein was shown to exchange GTP for bound GDP when activated by a hormone receptor. The GTP-carrying form then binds to activated adenylate cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

[0005] The membrane protein gene superfamily of G-protein coupled receptors has been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane α-helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuroreceptors.

[0006] G-protein coupled receptors (otherwise known as 7TM receptors) have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. The G-protein family of coupled receptors includes dopamine receptors which bind to neuroleptic drugs used for treating psychotic and neurological disorders. Other examples of members of this family include, but are not limited to, calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins, endothelial differentiation gene-1, rhodopsins, odorant, and cytomegalovirus receptors.

[0007] Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 has been implicated in signal transduction.

[0008] Phosphorylation and lipidation (palmitylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxy terminus. For several G-protein coupled receptors, such as the β-adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

[0009] For some receptors, the ligand binding sites of G-protein coupled receptors are believed to comprise hydrophilic sockets formed by several G-protein coupled receptor transmembrane domains, said sockets being surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form a polar ligand binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand binding site, such as the TM3 aspartate residue. TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

[0010] G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson et al.. Endoc. Rev., 1989, 10:317-331). Different G-protein α-subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors has been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors. G-protein coupled receptors are found in numerous

sites within a mammalian host. Over the past 15 years, nearly 350 therapeutic agents targeting 7 transmembrane (7 TM) receptors have been successfully introduced onto the market.

[0011] This indicates that these receptors have an established, proven history as therapeutic targets. Clearly there is a need for identification and characterization of further receptors which can play a role in preventing, ameliorating or correcting dysfunctions or diseases, including, but not limited to, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome.

SUMMARY OF THE INVENTION

[0012] In one aspect, the invention relates to HFIAO41 polypeptides and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such HFIAO41 polypeptides and polynucleotides. Such uses include the treatment of infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with HFIAO41 imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate HFIAO41 activity or levels.

DESCRIPTION OF THE INVENTION

Definitions

[0013] The following definitions are provided to facilitate understanding of certain terms used frequently herein.

[0014] "HFIAO41" refers, among others, to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, or an allelic variant thereof.

[0015] "Receptor Activity" or "Biological Activity of the Receptor" refers to the metabolic or physiologic function of said HFIAO41 including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said HFIAO41.

[0016] "HFIAO41 gene" refers to a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO: 1 or allelic variants thereof and/or their complements.

[0017] "Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

[0018] "Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

[0019] "Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single-and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

[0020] "Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly

referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62.

[0021] "Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

[0022] "Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCS program package (Devereux, J., *et al., Nucleic Acids Research* (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. *et al., J Molec Biol* (1990) 215:403).

[0023] As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence of SEQ ID NO: 1 is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence of SEQ ID NO: 1. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5 or 3 terminal positions of the reference nucleotide sequence

or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[0024] Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference amino acid sequence of SEQ ID NO:2 is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of SEQ ID NO: 2. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% ofthe amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Polypeptides of the Invention

[0025] In one aspect, the present invention relates to HFIAO41 polypeptides (or HFIAO41 proteins). The HFIAO41 polypeptides include the polypeptides of SEQ ID NOS:2 and 4; as well as polypeptides comprising the amino acid sequence of SEQ ID NO:2; and polypeptides comprising the amino acid sequence which have at least 87% identity to that of SEQ ID NO:2 over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO: 2. Furthermore, those with at least 97-99% are highly preferred. Also included within HFIAO41 polypeptides are polypeptides having the amino acid sequence which have at least 87% identity to the polypeptide having the amino acid sequence of SEQ ID NO: 2 over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO: 2. Furthermore, those with at least 97-99% are highly preferred. Preferably HFIAO41 polypeptides exhibit at least one biological activity of the receptor.

[0026] The HFIAO41 polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[0027] Fragments of the HFIAO41 polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence ofthe aforementioned HFIAO41 polypeptides. As with HFIAO41 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of HFIAO41 polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

[0028] Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of HFIAO41 polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate receptor activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

[0029] Preferably, all of these polypeptide fragments retain the biological activity of the receptor, including antigenic activity. Among the most preferred fragment is that having the amino acid sequence of SEQ ID NO: 4. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions -- i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination.

[0030] The HFIAO41 polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well

understood in the art.

Polynucleotides of the Invention

[0031] Another aspect of the invention relates to HFIAO41 polynucleotides. HFIAO41 polynucleotides include isolated polynucleotides which encode the HFIAO41 polypeptides and fragments, and polynucleotides closely related thereto. More specifically, HFIAO41 polynucleotide of the invention include a polynucleotide comprising the nucleotide sequence contained in SEQ ID NO: 1 encoding a HFIAO41 polypeptide of SEQ ID NO: 2, and polynucleotides having the particular sequences of SEQ ID NOS: 1 and 3. HFIAO41 polynucleotides further include a polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the HFIAO41 polypeptide of SEQ ID NO:2, and a polynucleotide comprising a nucleotide sequence that is at least 80% identical to that of SEQ ID NO: 1 over its entire length. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. Also included under HFIAO41 polynucleotides are a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO: 1 to hybridize under conditions useable for amplification or for use as a probe or marker. The invention also provides polynucleotides which are complementary to such HFIAO41 polynucleotides.

[0032] HFIAO41 of the invention is structurally related to other proteins of the G-protein coupled receptor family, as shown by the results of sequencing the cDNA encoding human HFIAO41. The cDNA sequence of SEQ ID NO: 1 contains an open reading frame (nucleotide number 249 to 1298) encoding a polypeptide of 350 amino acids of SEQ ID NO:2. The amino acid sequence of Table 1 (SEQ ID NO:2) has about 86% identity (using FASTA) in 350 amino acid residues with Bovine Possible Gustatory Receptor Type B (Biochem. Biophys. Res. Commun. 194(1) 504-511, 1993). Furthermore, HFIAO41 (SEQ ID NO: 2) is 39% identical to Human EBV Induced G-protein Coupled Receptor over 332 amino acid residues (J. Virol. 67(4) 2209-2220, 1993). The nucleotide sequence of Table 1 (SEQ ID NO: 1) has about 64% identity (using FASTA) in 2407 nucleotide residues with Bovine Possible Gustatory Receptor Type B (Biochem. Biophys. Res. Commun. 194(1) 504-511, 1993). Thus, HFIAO41 polypeptides and polynucleotides of the present invention are expected to have, inter alia, similar biological functions/properties to their homologous polypeptides and polynucleotides, and their utility is obvious to anyone skilled in the art.

Table 1[a]

1	GAACCGAGAT TATACCATTA CAGTCCAGCC TGGGCAACAG AGCCAGAGAC
51	CCTGTCATAA ATAAATAAAT AAACAAACAA ACAAATAAAA ATGGTGGAGT
101	cTGAAAAAGG ACTGGGTCAG CAAGAATAAA AACACAAAAC AGCTGGAGGA
151	GCCAAGATGG CCGAATAGGA ACAGCTCCGG TcTACAGCTC CCAGCGTGAG
201	CGACGCAGAA GACGGGTGAT TTCTGCATTT CCATCTGAGA TTGGAGCCAT
251	GGCTTTGGAA CAGAACCAGT CAACAGATTA TTATTATGAG GAAAATGAAA
301	TGAATGGCAC TTATGACTAC AGTCAATATG AACTGATCTG TATCAAAGAA
351	GATGTCAGAG AATTTGCAAA AGTTTTCCTC CCTGTATTCC TCACAATAGT
401	TTTCGTCATT GGACTTGCAG GCAATTCCAT GGTAGTGGCA ATTTATGCCT
451	ATTACAAGAA ACAGAGAACC AAAACAGATG TGTACATCCT GAATTTGGCT
501	GTAGCAGATT TACTCCTTCT ATTCACTCTG CCTTTTTGGG CTGTTAATGC
551	AGTTCATGGG TGGGTTTTAG GGAAAATAAT GTGCAAAATA ACTTCAGCCT
601	TGTACACACT AAACTTTGTC TCTGGAATGC AGTTTCTGGC TTGTATCAGC
651	ATAGACAGAT ATGTGGCAGT AACTAAAGTC CCCAGCCAAT CAGGAGTGGG

```
 701  AAAACCATGC TGGATCATCT GTTTCTGTGT CTGGATGGCT GCCATCTTGC
 751  TGAGCATACC CCAGCTGGTT TTTTATACAG TAAATGACAA TGCTAGGTGC
 801  ATTCCCATTT TCCCCCGCTA CCTAGGAACA TCAATGAAAG CATTGATTCA
 851  AATGCTAGAG ATCTGCATTG GATTTGTAGT ACCCTTTCTT ATTATGGGGG
 901  TGTGCTACTT TATCACAGCA AGGACACTCA TGAAGATGCC AAACATTAAA
 951  ATATCTCGAC CCCTAAAAGT TCTGCTCACA GTCGTTATAG TTTTCATTGT
1001  CACTCAACTG CCTTATAACA TTGTCAAGTT CTGCCGAGCC ATAGACATCA
1051  TCTACTCCCT GATCACCAGC TGCAACATGA GCAAACGCAT GGACATCGCC
1101  ATCCAAGTCA CAGAAAGCAT CGCACTCTTT CACAGCTGCC TCAACCCAAT
1151  CCTTTATGTT TTTATGGGAG CATCTTTCAA AAACTACGTT ATGAAAGTGG
1201  CCAAGAAATA TGGGTCCTGG AGAAGACAGA GACAAAGTGT GGAGGAGTTT
1251  CCTTTTGATT CTGAGGGTCC TACAGAGCCA ACCAGTACTT TTAGCATTTA
1301  AAGGTAAAAC TGCTCTGCCT TTTGCTTGGA TACATATGAA TGATGCTTTC
1351  CCCTCAAATA AAACATCTGC ATTATTCTGA AACTCAAATC TCAGACGCCG
1401  TGGTTGCAAC TTATAATAAA GAATGGGTTG GGGGAAGGGG GAGAAATAAA
1451  AGCCAAGAAG AGGAAACAAG ATAATAAATG TACAAAACAT GAAAATTAAA
1501  ATGAACAATA TAGGAAAATA ATTGTAACAG GCATAAGTGA ATAACACTCT
1551  GCTGTAACGA AGAAGAGCTT TGTGGTGATA ATTTTGTATC TTGGTTGCAG
1601  TGGTGCTTAT ACAAATCTAC ACAAGTGATA AAATGACACA GAACTATATA
1651  CACACATTGT ACCAATTTCA ATTTCCTGGT TTTGACATTA TAGTATAATT
```

```
1701  ATGTAAGATG GAACCATTGG GGAAAACTGG GTGAAGGGTA CCCAGGACCA

1751  CTCTGTACCA TCTTTGTAAC TTCCTGTGAA TTTATAATAA TTTCAAAATA

1801  AAACAAGTTA AAAAAAAACC CACTATGCTA TAAGTTAGGC CATCTAAAAC

1851  AGATTATTAA AGAGGTTCAT GTTAAAAGGC ATTTATAATT ATTTTTAATT

1901  ATCTAAGTTT TAATACAAGA ACGATTTCCT GCATAATTTT AGTACTTGAA

1951  TAAGTATGCA GCAGAACTCC AACTATCTTT TTTCCTGTTT TTTTTAAATT

2001  TGTAAGTAAT TTTATAAAAT CCACCTCCTC CAAAAAAGCA ATAAAAAAAA

2051  AACAAACTAT AATAAGCTTT TCTGATTCTT TTCAAAACAT TCCTGGTAAG

2101  TTCCTAAAGA CATAATTTGC TTCTATGATG TCAACTTTCT TACTAATAAC

2151  TGGTTATCAT GACAAATGTT AGGTTTATCA TATATAGTCT AGGTGTAATC

2201  CTCAGACTAT CATTTTCATC TGGGTTCCAA TTTCTTAACT TCCTAAAGAA

2251  TTCATCTGTT TATACAAGTC TACCACTGCC GATTGACTAA AAAATACATT

2301  ATCCCATGCA TAAAATGTCC TATTTTCATT TAAACACTTT ATTTTTGAGT

2351  AATAAAAATA TGTACCACAA TAAATTATTG TTAATTAACA AAAAAAAAAA

2401  AAAAAAA
```

[a] A nucleotide sequence of a human HFIAO41 (SEQ ID NO: 1).

Table 2[b]

```
   1  MALEQNQSTD YYYEENEMNG TYDYSQYELI CIKEDVREFA KVFLPVFLTI

  51  VFVIGLAGNS MVVAIYAYYK KQRTKTDVYI LNLAVADLLL LFTLPFWAVN

 101  AVHGWVLGKI MCKITSALYT LNFVSGMQFL ACISIDRYVA VTKVPSQSGV

 151  GKPCWIICFC VWMAAILLSI PQLVFYTVND NARCIPIFPR YLGTSMKALI
```

```
201  QMLEICIGFV VPFLIMGVCY FITARTLMKM PNIKISRPLK VLLTVVIVFI

251  VTQLPYNIVK FCRAIDIIYS LITSCNMSKR MDIAIQVTES IALFHSCLNP

301  ILYVFMGASF KNYVMKVAKK YGSWRRQRQS VEEFPFDSEG PTEPTSTFSI
```

An amino acid sequence of a human HFIAO41 (SEQ ID NO: 2).

[0033] One polynucleotide of the present invention encoding HFIAO41 may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in cells of human synovial fibroblasts, placenta using the expressed sequence tag (EST) analysis (Adams. M.D., *et al. Science* (1991) 252:1651-1656; Adams, M.D. *et al., Nature,* (1992) *355:632-634;* Adams, M.D., *et al., Nature* (1995) 377 Supp:3-174). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

[0034] The nucleotide sequence encoding HFIAO41 polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence contained in Table I (nucleotide number 249 to 1298 of SEQ ID NO: 1), or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2.

[0035] When the polynucleotides of the invention are used for the recombinant production of HFIAO41 polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al., Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain noncoding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

[0036] Further preferred embodiments are polynucleotides encoding HFIAO41 variants comprising the amino acid sequence of HFIAO41 polypeptide of Table 2 (SEQ ID NO:2) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination. Among the preferred polynucleotides of the present invention is contained in Table 3 (SEQ ID NO: 3) encoding the amino acid sequence of Table 4 (SEQ ID NO: 4).

Table 3[c]

GCTTTGGAACAGAACCAGTCAACAGATTATTATTATGAGGAAAATGAAATGAATGGCACT
TATGACTACAGTCAATATGAACTGATCTGTATCAAAGAAGATGTCAGAGAATTTGCAAAA
GTTTTCCTCCCTGTATTCCTCACAATAGTTTTCGTCATTGGACTTGCAGGCAATTCCATG
GTAGTGGCAATTTATGCCTATTACAAGAAACAGAGAACCAAAACAGATGTGTACATCCTG
AATTTGGCTGTAGCAGATTTACTCCTTCTATTCACTCTGCCTTTTTGGGCTGTTAATGCA
GTTCATGGGTGGGTTTTAGGGAAAATAATGTGCAAAATAACTTCAGCCTTGTACACACTA
AACTTTGTCTCTGGAATGCAGTTTCTGGCTTGTATCAGCATAGACAGATATGTGGCAGTA
ACTAAAGTCCCCAGCCAATCAGGAGTGGGAAAACCATGCTGGATCATCTGTTTCTGTGTC
TGGATGGCTGCCATCTTGCTGAGCATACCCCAGCTGGTTTTTTATACAGTAAATGACAAT
GCTAGGTGCATTCCCATTTTCCCCCGCTACCTAGGAACATCAATGAAAGCATTGATTCAA
ATGCTAGAGATCTGCATTGGATTTGTAGTACCCTTTCTTATTATGGGGGTGTGCTACTTT
ATCACAGCAAGGACACTCATGAAGATGCCAAACATTAAAATATCTCGACCCCTAAAAGTT
CTGCTCACAGTCGTTATAGTTTTCATTGTCACTCAACTGCCTTATAACATTGTCAAGTTC
TGCCGAGCCATAGACATCATCTACTCCCTGATCACCAGCTGCAACATGAGCAAACGCATG
GACATCGCCATCCAAGTCACAGAAAGCATCGCACTCTTTCACAGCTGCCTCAACCCAATC
CTTTATGTTTTTATGGGAGCATCTTTCAAAAACTACGTTATGAAAGTGGCCAAGAAATAT
GGGTCCTGGAGAAGACAGAGACAAAGTGTGGAGGAGTTTCCTTTTGATTCTGAGGGTCCT
ACAGAGCCAACCAGTACTTTTAGCATTTAAAGGTAAAACTGCTCTGCCTTTTGCTTGGAT
ACATATGAATGATGCTTTCCCCTCAAATAAAACATCTGCATTATTCTGAAACTCAAATCT
CAGACGCCGTGGTTGCAACTTATAATAAAGAATGGGTTGGGGGAAGGGGGAGAAATAAAA
GCCAAGAAGAGGAAACAAGATAATAAATGTACAAAACATGAAAATTAAAATGAACAATAT
AGGAAAATAATTGTAACAGGCATAAGTGAATAACACTCTGCTGTAACGAAGAAGAGCTTT
GTGGTGATAATTTTGTATCTTGGTTGCAGTGGTGCTTATACAAATCTACACAAGTGATAA
AATGACACAGAACTATATACACACATTGTACCAATTTCAATTTCCTGGTTTTGACATTAT
AGTATAATTATGTAAGATGGAACCATTGGGGAAAACTGGGTGAAGGGTACCCAGGACCAC
TCTGTACCATCTTTGTAACTTCCTGTGAATTTATAATAATTTCAAAATAAAACAAGTTAA
AAAAAAACCCACTATGCTATAAGTTAGGCCATCTAAAACAGATTATTAAAGAGGTTCATG
TTAAAAGGCATTTATAATTATTTTTAATTATCTAAGTTTTAATACAAGAACGATTTCCTG
CATAATTTTAGTACTTGAATAAGTATGCAGCAGAACTCCAACTATCTTTTTTCCTGTTTT
TTTTAAATTTGTAAGTAATTTTATAAAATCCACCTCCTCCAAAAAAGCAATAAAAAAAAA
ACAAACTATAATAAGCTTTTCTGATTCTTTTCAAAACATTCCTGGTAAGTTCCTAAAGAC
ATAATTTGCTTCTATGATGTCAACTTTCTTACTAATAACTGGTTATCATGACAAATGTTA
GGTTTATCATATATAGTCTAGGTGTAATCCTCAGACTATCATTTTCATCTGGGTTCCAAT
TTCTTAACTTCCTAAAGAATTCATCTGTTTATACAAGTCTACCACTGCCGATTGACTAAA
AAATACATTATCCCATGCATAAAATGTCCTATTTTCATTTAAACACTTTATTTTTGAGTA
ATAAAAATATGTACCACAATAAATTATTGTTAATTAACAAAAAAAAAAAAAAAAAA

[c] A partial nucleotide sequence of a human HFIAO41 (SEQ ID NO: 3).

Table 4[d]

ALEQNQSTDYYYEENEMNGTYDYSQYELICIKEDVREFAKVFLPVFLTIVFVIGLAGNSM
VVAIYAYYKKQRTKTDVYILNLAVADLLLLFTLPFWAVNAVHGWVLGKIMCKITSALYTL
NFVSGMQFLACISIDRYVAVTKVPSQSGVGKPCWIICFCVWMAAILLSIPQLVFYTVNDN
ARCIPIFPRYLGTSMKALIQMLEICIGFVVPFLIMGVCYFITARTLMKMPNIKISRPLKV
LLTVVIVFIVTQLPYNIVKFCRAIDIIYSLITSCNMSKRMDIAIQVTESIALFHSCLNPI
LYVFMGASFKNYVMKVAKKYGSWRRQRQSVEEFPFDSEGPTEPTSTFSI

[d] A partial amino acid sequence of a human HFIAO41 (SEQ ID NO: 4).

[0037] The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 80%, and preferably at least 90%, and more preferably at least 95%, yet even more preferably 97-99% identity between the sequences.

[0038] Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO: 1 or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding HFIAO41 and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs from species other than human) that have a high sequence similarity to the HFIAO41 gene. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that ofthe referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

[0039] In one embodiment, to obtain a polynucleotide encoding HFIAO41 polypeptide, including homologs and orthologs from species other than human, comprises the steps of screening an appropriate library under stingent hybridization conditions with a labeled probe having the SEQ ID NO: 1 or a fragment thereof (including that of SEQ ID NO: 3), and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or alternatively conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

[0040] The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

Vectors, Host Cells, Expression

[0041] The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

[0042] For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY* (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

[0043] Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli, Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

[0044] A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as bac-

uloviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al., MOLECULAR CLONING, A LABORATORY MANUAL* (*supra*).

[0045] For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

[0046] If the HFIAO41 polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If HFIAO41 polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

[0047] HFIAO41 polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Diagnostic Assays

[0048] This invention also relates to the use of HFIAO41 polynucleotides for use as diagnostic reagents. Detection of a mutated form of HFIAO41 gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of HFIAO41. Individuals carrying mutations in the HFIAO41 gene may be detected at the DNA level by a variety of techniques.

[0049] Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled HFIAO41 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers *et al., Science* (1985) 230:1242. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. See Cotton *et al., Proc Natl Acad Sci USA* (1985) 85: 4397-4401. In another embodiment, an array of oligonucleotides probes comprising HFIAO41 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability. (See for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996)).

[0050] The diagnostic assays offer a process for diagnosing or determining a susceptibility to infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome through detection of mutation in the HFIAO41 gene by the methods described.

[0051] In addition, infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of HFIAO41 polypeptide or HFIAO41 mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of

polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as an HFIAO41, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

[0052] Thus in another aspect, the present invention relates to a diagonostic kit for a disease or suspectability to a disease, particularly infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, which comprises:

(a) a HFIAO41 polynucleotide, preferably the nucleotide sequence of SEQ ID NO: 1, or a fragment thereof;
(b) a nucleotide sequence complementary to that of (a);
(c) a HFIAO41 polypeptide, preferably the polypeptide of SEQ ID NO: 2, or a fragment thereof; or
(d) an antibody to a HFIAO41 polypeptide, preferably to the polypeptide of SEQ ID NO: 2. It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Chromosome Assays

[0053] The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent ofthe disease.

Antibodies

[0054] The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the HFIAO41 polypeptides. The term "immunospecific" means that the antibodies have substantiall greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

[0055] Antibodies generated against the HFIAO41 polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al., Immunology Today* (1983)4:72) and the EBV-hybridoma technique (Cole *et al.,* MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985).

[0056] Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

[0057] The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

[0058] Antibodies against HFIAO41 polypeptides may also be employed to treat infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others.

Vaccines

[0059] Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with HFIAO41 polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome, among others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering HFIAO41 polypeptide via a vector directing expression of HFIAO41 polynucleotide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

[0060] Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a HFIAO41 polypeptide wherein the composition comprises a HFIAO41 polypeptide or HFIAO41 gene. The vaccine formulation may further comprise a suitable carrier. Since HFIAO41 polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain antioxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Screening Assays

[0061] The HFIAO41 polypeptide of the present invention may be employed in a screening process for compounds which bind the receptor and which activate (agonists) or inhibit activation of (antagonists) the receptor polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan *et al., Current Protocols in Immunology* 1(2):Chapter 5 (1991).

[0062] HFIAO41 polypeptides are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate HFIAO41 on the one hand and which can inhibit the function of HFIAO41 on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions as infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions as infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; pain; cancers; diabetes, obesity; anorexia; bulimia; asthma; Parkinson's disease; acute heart failure; hypotension; hypertension; urinary retention; osteoporosis; angina pectoris; myocardial infarction; ulcers; asthma; allergies; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles dela Tourett's syndrome.

[0063] In general, such screening procedures involve producing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof. Such cells include cells from mammals, yeast, *Drosophila* or *E. coli*. Cells expressing the receptor (or cell membrane containing the expressed receptor) are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response.

[0064] One screening technique includes the use of cells which express receptor of this invention (for example, transfected CHO cells) in a system which measures extracellular pH or intracellular calcium changes caused by receptor activation. In this technique, compounds may be contacted with cells expressing the receptor polypeptide of the present invention. A second messenger response, e.g., signal transduction, pH changes, or changes in calcium level, is then

measured to determine whether the potential compound activates or inhibits the receptor.

[0065] Another method involves screening for receptor inhibitors by determining inhibition or stimulation of receptor-mediated cAMP and/or adenylate cyclase accumulation. Such a method involves transfecting a eukaryotic cell with the receptor of this invention to express the receptor on the cell surface. The cell is then exposed to potential antagonists in the presence of the receptor of this invention. The amount of cAMP accumulation is then measured. If the potential antagonist binds the receptor, and thus inhibits receptor binding, the levels of receptor-mediated cAMP, or adenylate cyclase, activity will be reduced or increased. Another method for detecting agonists or antagonists for the receptor of the present invention is the yeast based technology as described in U.S. Patent No. 5,482,835.

[0066] The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the receptor is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the receptor, using detection systems appropriate to the cells bearing the receptor at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed.

[0067] Further, the assays may simply comprise the steps of mixing a candidate compound with a solution containing a HFIAO41 polypeptide to form a mixture, measuring HFIAO41 activity in the mixture, and comparing the HFIAO41 activity of the mixture to a standard.

[0068] The HFIAO41 cDNA, protein and antibodies to the protein may also be used to configure assays for detecting the effect of added compounds on the production of HFIAO41 mRNA and protein in cells. For example, an ELISA may be constructed for measuring secreted or cell associated levels of HFIAO41 protein using monoclonal and polyclonal antibodies by standard methods known in the art, and this can be used to discover agents which may inhibit or enhance the production of HFIAO41 (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues. Standard methods for conducting screening assays are well understood in the art.

[0069] Examples of potential HFIAO41 antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligand of the HFIAO41, e.g., a fragment ofthe ligand, or small molecules which bind to the receptor but do not elicit a response, so that the activity of the receptor is prevented.

[0070] Thus in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for HFIAO41 polypeptides; or compounds which decrease or enhance the production of HFIAO41 polypeptides, which comprises:

(a) a HFIAO41 polypeptide, preferably that of SEQ ID NO:2;
(b) a recombinant cell expressing a HFIAO41 polypeptide, preferably that of SEQ ID NO:2;
(c) a cell membrane expressing a HFIAO41 polypeptide; preferably that of SEQ ID NO: 2; or
(d) antibody to a HFIAO41 polypeptide, preferably that of SEQ ID NO: 2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Prophylactic and Therapeutic Methods

[0071] This invention provides methods of treating an abnormal conditions related to both an excess of and insufficient amounts of HFIAO41 activity.

[0072] If the activity of HFIAO41 is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the HFIAO41, or by inhibiting a second signal, and thereby alleviating the abnormal condition.

[0073] In another approach, soluble forms of HFIAO41 polypeptides still capable of binding the ligand in competition with endogenous HFIAO41 may be administered. Typical embodiments of such competitors comprise fragments of the HFIAO41 polypeptide.

[0074] In still another approach, expression of the gene encoding endogenous HFIAO41 can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. See, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Alternatively, oligonucleotides which form triple helices with the gene can be supplied. See, for example, Lee *et al., Nucleic Acids Res* (1979) 6:3073; Cooney et *al., Science* (1988) 241:456; Dervan *et al., Science* (1991) 251:1360. These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

[0075] For treating abnormal conditions related to an under-expression of HFIAO41 and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates HFIAO41, i.e., an agonist as described above, in combination with a pharmaceutically accept-

able carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of HFIAO41 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For overview of gene therapy, see Chapter 20, *Gene Therapy and other Molecular Genetic-based Therapeutic Approaches*, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Formulation and Administration

[0076] Peptides, such as the soluble form of HFIAO41 polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

[0077] Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

[0078] Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

[0079] The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 μg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

[0080] Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Example 1: Mammalian Cell Expression

[0081] The receptors of the present invention are expressed in either human embryonic kidney 293 (HEK293) cells or adherent dhfr CHO cells. To maximize receptor expression, typically all 5' and 3' untranslated regions (UTRs) are removed from the receptor cDNA prior to insertion into a pCDN or pCDNA3 vector. The cells are transfected with individual receptor cDNAs by lipofectin and selected in the presence of 400 mg/ml G418. After 3 weeks of selection, individual clones are picked and expanded for further analysis. HEK293 or CHO cells transfected with the vector alone serve as negative controls. To isolate cell lines stably expressing the individual receptors, about 24 clones are typically selected and analyzed by Northern blot analysis. Receptor mRNAs are generally detectable in about 50% of the G418-resistant clones analyzed.

Example 2 Ligand bank for binding and functional assays.

[0082] A bank of over 200 putative receptor ligands has been assembled for screening. The bank comprises: transmitters, hormones and chemokines known to act via a human seven transmembrane (7TM) receptor; naturally occurring compounds which may be putative agonists for a human 7TM receptor, non-mammalian, biologically active peptides for which a mammalian counterpart has not yet been identified; and compounds not found in nature, but which activate 7TM receptors with unknown natural ligands. This bank is used to initially screen the receptor for known ligands, using both functional (i.e . calcium, cAMP, microphysiometer, oocyte electrophysiology, etc, see below) as well as binding

assays.

Example 3: Ligand Binding Assays

[0083] Ligand binding assays provide a direct method for ascertaining receptor pharmacology and are adaptable to a high throughput format. The purified ligand for a receptor is radiolabeled to high specific activity (50-2000 Ci/mmol) for binding studies. A determination is then made that the process of radiolabeling does not diminish the activity of the ligand towards its receptor. Assay conditions for buffers, ions, pH and other modulators such as nucleotides are optimized to establish a workable signal to noise ratio for both membrane and whole cell receptor sources. For these assays, specific receptor binding is defined as total associated radioactivity minus the radioactivity measured in the presence of an excess of unlabeled competing ligand. Where possible, more than one competing ligand is used to define residual nonspecific binding.

Example 4: Functional Assay in Xenopus Oocytes

[0084] Capped RNA transcripts from linearized plasmid templates encoding the receptor cDNAs of the invention are synthesized in vitro with RNA polymerases in accordance with standard procedures. In vitro transcripts are suspended in water at a final concentration of 0.2 mg/ml. Ovarian lobes are removed from adult female toads, Stage V defolliculated oocytes are obtained, and RNA transcripts (10 ng/oocyte) are injected in a 50 nl bolus using a microinjection apparatus. Two electrode voltage clamps are used to measure the currents from individual Xenopus oocytes in response to agonist exposure. Recordings are made in Ca2+ free Barth's medium at room temperature. The Xenopus system can be used to screen known ligands and tissue/cell extracts for activating ligands.

Example 5: Microphysiometric Assays

[0085] Activation of a wide variety of secondary messenger systems results in extrusion of small amounts of acid from a cell. The acid formed is largely as a result of the increased metabolic activity required to fuel the intracellular signaling process. The pH changes in the media surrounding the cell are very small but are detectable by the CYTOSENSOR microphysiometer (Molecular Devices Ltd., Menlo Park, CA). The CYTOSENSOR is thus capable of detecting the activation of a receptor which is coupled to an energy utilizing intracellular signaling pathway such as the G-protein coupled receptor of the present invention.

Example 6: Extract/Cell Supernatant Screening

[0086] A large number of mammalian receptors exist for which there remains, as yet, no cognate activating ligand (agonist). Thus, active ligands for these receptors may not be included within the ligands banks as identified to date. Accordingly, the 7TM receptor of the invention is also functionally screened (using calcium, cAMP, microphysiometer, oocyte electrophysiology, etc., functional screens) against tissue extracts to identify natural ligands. Extracts that produce positive functional responses can be sequentially subfractionated until an activating ligand is isolated and identified.

Example 7: Calcium and cAMP Functional Assays

[0087] 7TM receptors which are expressed in HEK 293 cells have been shown to be coupled functionally to activation of PLC and calcium mobilization and/or cAMP stimulation or inhibition. Basal calcium levels in the HEK 293 cells in receptor-transfected or vector control cells were observed to be in the normal, 100 nM to 200 nM, range. HEK 293 cells expressing recombinant receptors are loaded with fura 2 and in a single day > 150 selected ligands or tissue/cell extracts are evaluated for agonist induced calcium mobilization. Similarly, HEK 293 cells expressing recombinant receptors are evaluated for the stimulation or inhibition of cAMP production using standard cAMP quantitation assays. Agonists presenting a calcium transient or cAMP fluctuation are tested in vector control cells to determine if the response is unique to the transfected cells expressing receptor.

[0088] All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

Annex to the description

[0089]

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: SmithKline Beecham Corporation

(ii) TITLE OF THE INVENTION: THE G-PROTEIN COUPLED RECEPTOR HFIAO41

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: F J Cleveland & Company
(B) STREET: 40/43 Chancery Lane
(C) CITY: London
(D) COUNTY:
(E) COUNTRY: United Kingdom
(F) POST CODE: WC2A 1JQ

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: TO BE ASSIGNED
(B) FILING DATE: 24-OCT-1997
(C) CLASSIFICATION: UNKNOWN

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/055,895
(B) FILING DATE: 15-AUG-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: CRUMP, Julian Richard John
(B) GENERAL AUTHORISATION NUMBER: 37127
(C) REFERENCE/DOCKET NUMBER: GH-70225

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: +44 171 504 5875
(B) TELEFAX: +44 171 831 .0749
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2407 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAACCGAGAT TATACCATTA CAGTCCAGCC TGGGCAACAG AGCCAGAGAC CCTGTCATAA 60
ATAAATAAAT AAACAAACAA ACAAATAAAA ATGGTGGAGT CTGAAAAAGG ACTGGGTCAG 120
CAAGAATAAA AACACAAAAC AGCTGGAGGA GCCAAGATGG CCGAATAGGA ACAGCTCCGG 180
TCTACAGCTC CCAGCGTGAG CGACGCAGAA GACGGGTGAT TTCTGCATTT CCATCTGAGA 240
TTGGAGCCAT GGCTTTGGAA CAGAACCAGT CAACAGATTA TTATTATGAG GAAAATGAAA 300
TGAATGGCAC TTATGACTAC AGTCAATATG AACTGATCTG TATCAAAGAA GATGTCAGAG 360
AATTTGCAAA AGTTTTCCTC CCTGTATTCC TCACAATAGT TTTCGTCATT GGACTTGCAG 420
GCAATTCCAT GGTAGTGGCA ATTTATGCCT ATTACAAGAA ACAGAGAACC AAAACAGATG 480
TGTACATCCT GAATTTGGCT GTAGCAGATT TACTCCTTCT ATTCACTCTG CCTTTTTGGG 540
CTGTTAATGC AGTTCATGGG TGGGTTTTAG GGAAAATAAT GTGCAAAATA ACTTCAGCCT 600
TGTACACACT AAACTTTGTC TCTGGAATGC AGTTTCTGGC TTGTATCAGC ATAGACAGAT 660
ATGTGGCAGT AACTAAAGTC CCCAGCCAAT CAGGAGTGGG AAAACCATGC TGGATCATCT 720
GTTTCTGTGT CTGGATGGCT GCCATCTTGC TGAGCATACC CCAGCTGGTT TTTTATACAG 780
TAAATGACAA TGCTAGGTGC ATTCCCATTT TCCCCCGCTA CCTAGGAACA TCAATGAAAG 840
CATTGATTCA AATGCTAGAG ATCTGCATTG GATTTGTAGT ACCCTTTCTT ATTATGGGGG 900
TGTGCTACTT TATCACAGCA AGGACACTCA TGAAGATGCC AAACATTAAA ATATCTCGAC 960
CCCTAAAAGT TCTGCTCACA GTCGTTATAG TTTTCATTGT CACTCAACTG CCTTATAACA 1020
TTGTCAAGTT CTGCCGAGCC ATAGACATCA TCTACTCCCT GATCACCAGC TGCAACATGA 1080
GCAAACGCAT GGACATCGCC ATCCAAGTCA CAGAAAGCAT CGCACTCTTT CACAGCTGCC 1140
TCAACCCAAT CCTTTATGTT TTTATGGGAG CATCTTTCAA AAACTACGTT ATGAAAGTGG 1200
CCAAGAAATA TGGGTCCTGG AGAAGACAGA GACAAAGTGT GGAGGAGTTT CCTTTTGATT 1260
CTGAGGGTCC TACAGAGCCA ACCAGTACTT TTAGCATTTA AAGGTAAAAC TGCTCTGCCT 1320
TTTGCTTGGA TACATATGAA TGATGCTTTC CCCTCAAATA AAACATCTGC ATTATTCTGA 1380
AACTCAAATC TCAGACGCCG TGGTTGCAAC TTATAATAAA GAATGGGTTG GGGGAAGGGG 1440
GAGAAATAAA AGCCAAGAAG AGGAAACAAG ATAATAAATG TACAAAACAT GAAAATTAAA 1500

```
ATGAACAATA TAGGAAAATA ATTGTAACAG GCATAAGTGA ATAACACTCT GCTGTAACGA   1560
AGAAGAGCTT TGTGGTGATA ATTTTGTATC TTGGTTGCAG TGGTGCTTAT ACAAATCTAC   1620
ACAAGTGATA AAATGACACA GAACTATATA CACACATTGT ACCAATTTCA ATTTCCTGGT   1680
TTTGACATTA TAGTATAATT ATGTAAGATG GAACCATTGG GGAAAACTGG GTGAAGGGTA   1740
CCCAGGACCA CTCTGTACCA TCTTTGTAAC TTCCTGTGAA TTTATAATAA TTTCAAAATA   1800
AAACAAGTTA AAAAAAAACC CACTATGCTA TAAGTTAGGC CATCTAAAAC AGATTATTAA   1860
AGAGGTTCAT GTTAAAAGGC ATTTATAATT ATTTTTAATT ATCTAAGTTT TAATACAAGA   1920
ACGATTTCCT GCATAATTTT AGTACTTGAA TAAGTATGCA GCAGAACTCC AACTATCTTT   1980
TTTCCTGTTT TTTTTAAATT TGTAAGTAAT TTTATAAAAT CCACCTCCTC CAAAAAAGCA   2040
ATAAAAAAAA AACAAACTAT AATAAGCTTT TCTGATTCTT TTCAAAACAT TCCTGGTAAG   2100
TTCCTAAAGA CATAATTTGC TTCTATGATG TCAACTTTCT TACTAATAAC TGGTTATCAT   2160
GACAAATGTT AGGTTTATCA TATATAGTCT AGGTGTAATC CTCAGACTAT CATTTTCATC   2220
TGGGTTCCAA TTTCTTAACT TCCTAAAGAA TTCATCTGTT TATACAAGTC TACCACTGCC   2280
GATTGACTAA AAAATACATT ATCCCATGCA TAAAATGTCC TATTTTCATT TAAACACTTT   2340
ATTTTTGAGT AATAAAAATA TGTACCACAA TAAATTATTG TTAATTAACA AAAAAAAAAA   2400
AAAAAAA                                                            2407
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Leu Glu Gln Asn Gln Ser Thr Asp Tyr Tyr Tyr Glu Glu Asn
 1               5                  10                  15
Glu Met Asn Gly Thr Tyr Asp Tyr Ser Gln Tyr Glu Leu Ile Cys Ile
            20                  25                  30
Lys Glu Asp Val Arg Glu Phe Ala Lys Val Phe Leu Pro Val Phe Leu
        35                  40                  45
Thr Ile Val Phe Val Ile Gly Leu Ala Gly Asn Ser Met Val Val Ala
    50                  55                  60
Ile Tyr Ala Tyr Tyr Lys Lys Gln Arg Thr Lys Thr Asp Val Tyr Ile
65                  70                  75                  80
Leu Asn Leu Ala Val Ala Asp Leu Leu Leu Phe Thr Leu Pro Phe
                85                  90                  95
Trp Ala Val Asn Ala Val His Gly Trp Val Leu Gly Lys Ile Met Cys
            100                 105                 110
```

```
Lys Ile Thr Ser Ala Leu Tyr Thr Leu Asn Phe Val Ser Gly Met Gln
        115                 120                 125
Phe Leu Ala Cys Ile Ser Ile Asp Arg Tyr Val Ala Val Thr Lys Val
    130                 135                 140
Pro Ser Gln Ser Gly Val Gly Lys Pro Cys Trp Ile Ile Cys Phe Cys
145                 150                 155                 160
Val Trp Met Ala Ala Ile Leu Leu Ser Ile Pro Gln Leu Val Phe Tyr
                165                 170                 175
Thr Val Asn Asp Asn Ala Arg Cys Ile Pro Ile Phe Pro Arg Tyr Leu
            180                 185                 190
Gly Thr Ser Met Lys Ala Leu Ile Gln Met Leu Glu Ile Cys Ile Gly
        195                 200                 205
Phe Val Val Pro Phe Leu Ile Met Gly Val Cys Tyr Phe Ile Thr Ala
    210                 215                 220
Arg Thr Leu Met Lys Met Pro Asn Ile Lys Ile Ser Arg Pro Leu Lys
225                 230                 235                 240
Val Leu Leu Thr Val Val Ile Val Phe Ile Val Thr Gln Leu Pro Tyr
                245                 250                 255
Asn Ile Val Lys Phe Cys Arg Ala Ile Asp Ile Ile Tyr Ser Leu Ile
            260                 265                 270
Thr Ser Cys Asn Met Ser Lys Arg Met Asp Ile Ala Ile Gln Val Thr
        275                 280                 285
Glu Ser Ile Ala Leu Phe His Ser Cys Leu Asn Pro Ile Leu Tyr Val
    290                 295                 300
Phe Met Gly Ala Ser Phe Lys Asn Tyr Val Met Lys Val Ala Lys Lys
305                 310                 315                 320
Tyr Gly Ser Trp Arg Arg Gln Arg Gln Ser Val Glu Glu Phe Pro Phe
                325                 330                 335
Asp Ser Glu Gly Pro Thr Glu Pro Thr Ser Thr Phe Ser Ile
            340                 345                 350
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2156 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCTTTGGAAC AGAACCAGTC AACAGATTAT TATTATGAGG AAAATGAAAT GAATGGCACT 60
TATGACTACA GTCAATATGA ACTGATCTGT ATCAAAGAAG ATGTCAGAGA ATTTGCAAAA 120
GTTTTCCTCC CTGTATTCCT CACAATAGTT TTCGTCATTG GACTTGCAGG CAATTCCATG 180
GTAGTGGCAA TTTATGCCTA TTACAAGAAA CAGAGAACCA AAACAGATGT GTACATCCTG 240
AATTTGGCTG TAGCAGATTT ACTCCTTCTA TTCACTCTGC CTTTTTGGGC TGTTAATGCA 300
GTTCATGGGT GGGTTTTAGG GAAAATAATG TGCAAAATAA CTTCAGCCTT GTACACACTA 360
AACTTTGTCT CTGGAATGCA GTTTCTGGCT TGTATCAGCA TAGACAGATA TGTGGCAGTA 420
ACTAAAGTCC CCAGCCAATC AGGAGTGGGA AAACCATGCT GGATCATCTG TTTCTGTGTC 480
TGGATGGCTG CCATCTTGCT GAGCATACCC CAGCTGGTTT TTTATACAGT AAATGACAAT 540
GCTAGGTGCA TTCCCATTTT CCCCCGCTAC CTAGGAACAT CAATGAAAGC ATTGATTCAA 600
ATGCTAGAGA TCTGCATTGG ATTTGTAGTA CCCTTTCTTA TTATGGGGGT GTGCTACTTT 660
ATCACAGCAA GGACACTCAT GAAGATGCCA AACATTAAAA TATCTCGACC CCTAAAAGTT 720
CTGCTCACAG TCGTTATAGT TTTCATTGTC ACTCAACTGC CTTATAACAT TGTCAAGTTC 780
TGCCGAGCCA TAGACATCAT CTACTCCCTG ATCACCAGCT GCAACATGAG CAAACGCATG 840
GACATCGCCA TCCAAGTCAC AGAAAGCATC GCACTCTTTC ACAGCTGCCT CAACCCAATC 900
CTTTATGTTT TTATGGGAGC ATCTTTCAAA AACTACGTTA TGAAAGTGGC CAAGAAATAT 960
GGGTCCTGGA GAAGACAGAG ACAAAGTGTG GAGGAGTTTC CTTTTGATTC TGAGGGTCCT 1020
ACAGAGCCAA CCAGTACTTT TAGCATTTAA AGGTAAAACT GCTCTGCCTT TTGCTTGGAT 1080
ACATATGAAT GATGCTTTCC CCTCAAATAA AACATCTGCA TTATTCTGAA ACTCAAATCT 1140
CAGACGCCGT GGTTGCAACT TATAATAAAG AATGGGTTGG GGGAAGGGGG AGAAATAAAA 1200
GCCAAGAAGA GGAAACAAGA TAATAAATGT ACAAAACATG AAAATTAAAA TGAACAATAT 1260
AGGAAAATAA TTGTAACAGG CATAAGTGAA TAACACTCTG CTGTAACGAA GAAGAGCTTT 1320
GTGGTGATAA TTTTGTATCT TGGTTGCAGT GGTGCTTATA CAAATCTACA CAAGTGATAA 1380
AATGACACAG AACTATATAC ACACATTGTA CCAATTTCAA TTTCCTGGTT TTGACATTAT 1440
AGTATAATTA TGTAAGATGG AACCATTGGG GAAAACTGGG TGAAGGGTAC CCAGGACCAC 1500
TCTGTACCAT CTTTGTAACT TCCTGTGAAT TTATAATAAT TTCAAAATAA AACAAGTTAA 1560
AAAAAAACCC ACTATGCTAT AAGTTAGGCC ATCTAAAACA GATTATTAAA GAGGTTCATG 1620
TTAAAAGGCA TTTATAATTA TTTTTAATTA TCTAAGTTTT AATACAAGAA CGATTTCCTG 1680
CATAATTTTA GTACTTGAAT AAGTATGCAG CAGAACTCCA ACTATCTTTT TTCCTGTTTT 1740
TTTTAAATTT GTAAGTAATT TTATAAAATC CACCTCCTCC AAAAAAGCAA TAAAAAAAAA 1800
ACAAACTATA ATAAGCTTTT CTGATTCTTT TCAAAACATT CCTGGTAAGT TCCTAAAGAC 1860
ATAATTTGCT TCTATGATGT CAACTTTCTT ACTAATAACT GGTTATCATG ACAAATGTTA 1920
GGTTTATCAT ATATAGTCTA GGTGTAATCC TCAGACTATC ATTTTCATCT GGGTTCCAAT 1980
TTCTTAACTT CCTAAAGAAT TCATCTGTTT ATACAAGTCT ACCACTGCCG ATTGACTAAA 2040
AAATACATTA TCCCATGCAT AAAATGTCCT ATTTTCATTT AAACACTTTA TTTTTGAGTA 2100
ATAAAAATAT GTACCACAAT AAATTATTGT TAATTAACAA AAAAAAAAAA AAAAAA 2156

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 349 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ala Leu Glu Gln Asn Gln Ser Thr Asp Tyr Tyr Tyr Glu Glu Asn Glu
 1               5                  10                  15
Met Asn Gly Thr Tyr Asp Tyr Ser Gln Tyr Glu Leu Ile Cys Ile Lys
            20                  25                  30
Glu Asp Val Arg Glu Phe Ala Lys Val Phe Leu Pro Val Phe Leu Thr
        35                  40                  45
Ile Val Phe Val Ile Gly Leu Ala Gly Asn Ser Met Val Val Ala Ile
    50                  55                  60
Tyr Ala Tyr Tyr Lys Lys Gln Arg Thr Lys Thr Asp Val Tyr Ile Leu
65                  70                  75                  80
Asn Leu Ala Val Ala Asp Leu Leu Leu Leu Phe Thr Leu Pro Phe Trp
                85                  90                  95
Ala Val Asn Ala Val His Gly Trp Val Leu Gly Lys Ile Met Cys Lys
            100                 105                 110
Ile Thr Ser Ala Leu Tyr Thr Leu Asn Phe Val Ser Gly Met Gln Phe
        115                 120                 125
Leu Ala Cys Ile Ser Ile Asp Arg Tyr Val Ala Val Thr Lys Val Pro
    130                 135                 140
Ser Gln Ser Gly Val Gly Lys Pro Cys Trp Ile Ile Cys Phe Cys Val
145                 150                 155                 160
Trp Met Ala Ala Ile Leu Leu Ser Ile Pro Gln Leu Val Phe Tyr Thr
                165                 170                 175
Val Asn Asp Asn Ala Arg Cys Ile Pro Ile Phe Pro Arg Tyr Leu Gly
            180                 185                 190
Thr Ser Met Lys Ala Leu Ile Gln Met Leu Glu Ile Cys Ile Gly Phe
        195                 200                 205
Val Val Pro Phe Leu Ile Met Gly Val Cys Tyr Phe Ile Thr Ala Arg
    210                 215                 220
Thr Leu Met Lys Met Pro Asn Ile Lys Ile Ser Arg Pro Leu Lys Val
225                 230                 235                 240
Leu Leu Thr Val Val Ile Val Phe Ile Val Thr Gln Leu Pro Tyr Asn
                245                 250                 255
Ile Val Lys Phe Cys Arg Ala Ile Asp Ile Ile Tyr Ser Leu Ile Thr
            260                 265                 270
Ser Cys Asn Met Ser Lys Arg Met Asp Ile Ala Ile Gln Val Thr Glu
        275                 280                 285
```

```
Ser Ile Ala Leu Phe His Ser Cys Leu Asn Pro Ile Leu Tyr Val Phe
    290                 295                 300
Met Gly Ala Ser Phe Lys Asn Tyr Val Met Lys Val Ala Lys Lys Tyr
305                 310                 315                 320
Gly Ser Trp Arg Arg Gln Arg Gln Ser Val Glu Glu Phe Pro Phe Asp
                325                 330                 335
Ser Glu Gly Pro Thr Glu Pro Thr Ser Thr Phe Ser Ile
            340                 345
```

Claims

1. An isolated polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the HFIAO41 polypeptide of SEQ ID NO:2; or a nucleotide sequence complementary to said isolated polynucleotide.

2. The polynucleotide of claim 1 wherein said polynucleotide comprises the nucleotide sequence contained in SEQ ID NO: 1 encoding the HFIAO41 polypeptide of SEQ ID NO2.

3. The polynucleotide of claim 1 wherein said polynucleotide comprises a nucleotide sequence that is at least 80% identical to that of SEQ ID NO: 1 over its entire length.

4. The polynucleotide of claim 3 which is polynucleotide of SEQ ID NO: 1.

5. The polynucleotide of claim 1 which is DNA or RNA.

6. A DNA or RNA molecule comprising an expression system, wherein said expression system is capable of producing a HFIAO41 polypeptide comprising an amino acid sequence, which has at least 87% identity with the polypeptide of SEQ ID NO:2 when said expression system is present in a compatible host cell.

7. A host cell comprising the expression system of claim 6.

8. A process for producing a HFIAO41 polypeptide comprising culturing a host of claim 7 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture.

9. A process for producing a cell which produces a HFIAO41 polypeptide thereof comprising transforming or transfecting a host cell with the expression system of claim 6 such that the host cell, under appropriate culture conditions, produces a HFIAO41 polypeptide.

10. A HFIAO41 polypeptide comprising an amino acid sequence which is at least 87% identical to the amino acid sequence of SEQ ID NO:2 over its entire length.

11. The polypeptide of claim 10 which comprises the amino acid sequence of SEQ ID NO:2.

12. An antibody immunospecific for the HFIAO41 polypeptide of claim 10.

13. A method for the treatment of a subject in need of enhanced activity or expression of HFIAO41 polypeptide of claim 10 comprising:

    (a) administering to the subject a therapeutically effective amount of an agonist to said receptor; and/or
    (b) providing to the subject an isolated polynucleotide comprising a nucleotide sequence that has at least 80%

identity to a nucleotide sequence encoding the HFIAO41 polypeptide of SEQ ID NO:2 over its entire length; or a nucleotide sequence complementary to said nucleotide sequence in a form so as to effect production of said receptor activity *in vivo*.

14. A method for the treatment of a subject having need to inhibit activity or expression of HFIAO41 polypeptide of claim 10 comprising:

(a) administering to the subject a therapeutically effective amount of an antagonist to said receptor; and/or
(b) administering to the subject a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said receptor; and/or
(c) administering to the subject a therapeutically effective amount of a polypeptide that competes with said receptor for its ligand.

15. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of HFIAO41 polypeptide of claim 10 in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said HFIAO41 polypeptide in the genome of said subject; and/or
(b) analyzing for the presence or amount of the HFIAO41 polypeptide expression in a sample derived from said subject.

16. A method for identifying agonists to HFIAO41 polypeptide of claim 10 comprising:

(a) contacting a cell which produces a HFIAO41 polypeptide with a candidate compound; and
(b) determining whether the candidate compound effects a signal generated by activation of the HFIAO41 polypeptide.

17. An agonist identified by the method of claim 16.

18. The method for identifying antagonists to HFIAO41 polypeptide of claim 10 comprising:

(a) contacting a cell which produces a HFIAO41 polypeptide with an agonist; and
(b) determining whether the signal generated by said agonist is diminished in the presence of a candidate compound.

19. An antagonist identified by the method of claim 18.

20. A recombinant host cell produced by a method of Claim 9 or a membrane thereof expressing a HFIAO41 polypeptide.